

PENERAPAN DATA MINING UNTUK MEMPREDIKSI KRITERIA NASABAH KREDIT

Angga Ginanjar Maburr^[1], Riani Lubis^[2]

^{1,2}Program Studi Teknik Informatika
Fakultas Teknik dan Ilmu Komputer Universitas Komputer Indonesia
Jl. Dipati Ukur No. 112-116 Bandung
Email : e.ngga@yahoo.com^[1], riani_lubis@yahoo.com^[2]

ABSTRAK

Salah satu tugas dari bidang *marketing* perbankan adalah melakukan analisis data nasabahnya untuk mengetahui nasabah-nasabah yang berpotensi melakukan kredit. Pada umumnya, metode yang digunakan untuk menganalisis data nasabah adalah dengan cara mengklasifikasikan semua nasabah yang telah melunasi angsuran kreditnya ke dalam target pemasaran, sehingga metode ini menyebabkan tingginya biaya operasional *marketing*. Oleh karena itu, penelitian ini dilakukan untuk membantu menyelesaikan permasalahan tersebut dengan merancang sebuah aplikasi *data mining* yang berfungsi untuk memprediksi kriteria nasabah kredit yang berpotensi melakukan peminjaman (kredit) terhadap bank. Bagian Dana Bank XY yang berlokasi di Kabupaten Bandung merupakan tempat yang dipilih oleh peneliti sebagai studi kasus, dengan asumsi Bagian Dana Bank XY tersebut telah mengalami permasalahan yang sama seperti yang telah dijelaskan di atas.

Teknik *data mining* yang diterapkan pada aplikasi yang dibangun adalah klasifikasi sedangkan metode klasifikasi yang digunakan adalah *Decision Tree* (pohon keputusan). Algoritma yang dipakai sebagai algoritma pembentuk pohon keputusannya adalah Algoritma C4.5. Adapun data yang di proses dalam penelitian ini adalah data angsuran nasabah kredit Bank XY pada bulan Juni 2009 dalam format Microsoft Excel.

Hasil penelitian ini adalah sebuah aplikasi yang dapat memudahkan Bagian Dana Bank XY dalam memperoleh target pemasaran kredit di masa mendatang.

Kata Kunci : Data mining, Marketing, Prediksi, Decision Tree, Algoritma C4.5.

1. PENDAHULUAN

Tugas Bagian Dana Bank XY selama ini adalah mencari nasabah dan/atau menghimpun dana

masyarakat berupa kredit, dana tabungan, deposito, simpanan, dan pinjaman lainnya. Metode yang digunakan untuk menentukan target pasar selama ini adalah mengklasifikasikan semua nasabah yang telah melunasi angsuran kreditnya ke dalam target pemasaran, padahal belum tentu semua nasabah tersebut akan tertarik melakukan peminjaman kredit kembali.

Oleh karena itu, untuk meningkatkan loyalitas nasabah bank tersebut, maka bagian ini dituntut untuk dapat mengambil keputusan yang tepat dalam menentukan strategi pemasaran, baik produk maupun layanan jasa perbankan yang akan ditawarkan kepada nasabah khususnya dalam masalah perkreditan. Hal ini dapat direalisasikan jika diperoleh informasi yang cukup untuk membantu proses pengambilan keputusan di bidang *marketing*, salah satunya adalah dengan melakukan peramalan atau memprediksi kriteria nasabah untuk kepentingan target pemasaran.

Data mining dapat digunakan oleh *marketer* untuk menganalisis pasar sehingga seorang *marketer* dapat memahami pasar yang sudah ada ataupun menemukan peluang-peluang yang baru untuk meningkatkan keuntungan termasuk memprediksi kriteria target pasar. Oleh karena itu, penelitian ini fokus pada pemanfaatan *data mining* untuk memprediksi kriteria nasabah kredit, sehingga dapat diketahui apakah nasabah yang bersangkutan merupakan nasabah yang berpotensi menjadi nasabah kredit yang produktif atau tidak.

Batasan masalah dari penelitian ini diantaranya adalah :

1. Membangun aplikasi *data mining* di Bagian Dana Bank XY Kabupaten Bandung.
2. Metode *data mining* yang digunakan adalah *decision tree* dengan algoritma *decision tree* yang digunakan adalah Algoritma C4.5
3. Data yang digunakan adalah data angsuran nasabah kredit Bank XY Kabupaten Bandung selama periode Juni 2009, dengan jumlah data yang digunakan sebanyak 250 *record*.
4. *Preprocessing* yang dilakukan terlebih dahulu terhadap data yang digunakan adalah *data integration and transformation, data cleaning,*

dan *data reduction* sehingga data tersebut siap untuk di-mining.

5. Proses *data integration and transformation, data cleaning*, dan *data reduction* tidak ditangani sistem.
6. Hasil dari proses *mining* adalah terbentuknya pola data berupa *rules* (aturan prediksi)
7. Pemodelan analisis pembangunan perangkat lunak yang digunakan adalah metode analisis berorientasi objek.

2. TINJAUAN PUSTAKA

Data mining adalah proses menganalisa data dari perspektif yang berbeda dan menyimpulkannya menjadi informasi-informasi penting yang dapat dipakai untuk meningkatkan keuntungan, memperkecil biaya pengeluaran, atau bahkan keduanya. Secara teknis, *data mining* dapat disebut sebagai proses untuk menemukan korelasi atau pola dari ratusan atau ribuan *field* dari sebuah relasional *database* yang besar. [2]

Kemampuan *Data mining* untuk mencari informasi bisnis yang berharga dari basis data yang sangat besar, dapat dianalogikan dengan penambangan logam mulia dari lahan sumbernya, teknologi ini dipakai untuk [2] :

1. Prediksi *trend* dan sifat-sifat bisnis, dimana *data mining* mengotomatisasi proses pencarian informasi prediksi di dalam basis data yang besar.
2. Penemuan pola-pola yang tidak diketahui sebelumnya, dimana *data mining* “menyapu” basis data, kemudian mengidentifikasi pola-pola yang sebelumnya tersembunyi dalam satu sapuan.

Dalam aplikasinya, *data mining* sebenarnya merupakan salah satu bagian proses *Knowledge Discovery in Database* (KDD) yang bertugas untuk mengekstrak pola atau model dari data dengan menggunakan suatu algoritma yang spesifik. Adapun proses KDD sebagai berikut : [1]

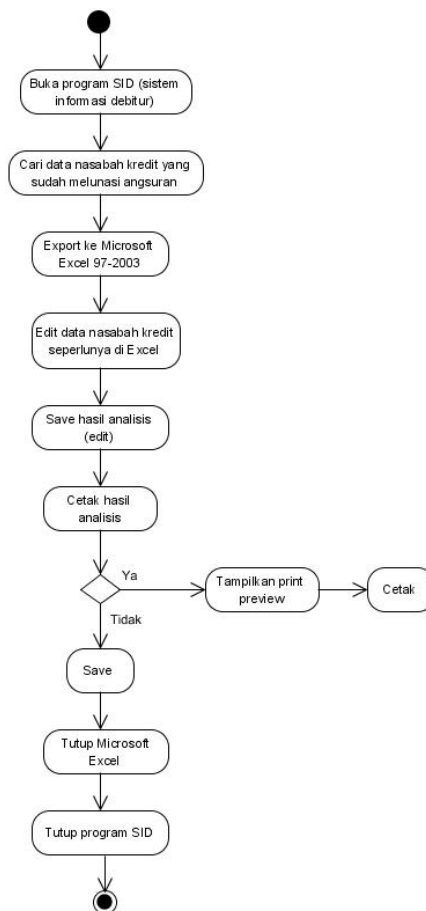
1. *Data Selection* : pemilihan data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai.
2. *Preprocessing* : sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* dengan tujuan untuk membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses *enrichment*, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.
3. *Transformation* : yaitu proses *coding* pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses *coding* dalam KDD merupakan proses kreatif dan sangat

tergantungan pada jenis atau pola informasi yang akan dicari dalam database.

4. *Data mining* : proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu.
5. *Interpretation / Evaluation* : pola informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut dengan *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya atau tidak.

3. ANALISIS SISTEM

Prosedur pencarian nasabah yang dilakukan oleh Bagian Dana Bank XY Kabupaten Bandung, dapat dilihat pada Gambar 1. Berdasarkan hasil observasi yang telah dilakukan, data nasabah kredit Bank XY Kabupaten Bandung tersimpan dalam *database microsoft excel* dengan nama MASTER_KREDIT hasil *import* dari Sistem Informasi Debitur (SID), dimana tabel tersebut terdiri dari 27 *fields* (Tabel 1).



Gambar 1. Activity Diagram Prosedur Pencarian Nasabah

Tabel 1
Struktur tabel MASTER_KREDIT

| |
|---------------------|
| Field |
| NASABAH_ID |
| NAMA_NASABAH |
| ALAMAT |
| JENIS_KELAMIN |
| TEMPATLAHIR |
| TGLLAHIR |
| UMUR |
| JENIS_ID |
| KODE_AGAMA |
| DESA |
| KECAMATAN |
| KOTA_KAB |
| PROPINSI |
| VERIFIKASI |
| TGL_REGISTER |
| KODEPOS |
| KODE_KANTOR |
| STATUS_MARITAL |
| BI_JENIS_PENGGUNAAN |
| BI_SEKTOR_EKONOMI |
| JML_PINJAMAN |
| NILAI_PINJAMAN |
| JML_ANGSURAN |
| KODE_INTEGRASI |
| STATUS |
| JENIS_AGUNAN |
| CEK |

Pada teknik classification, keluaran dari setiap data yang dijadikan parameter target atau class harus berupa bilangan bulat atau diskrit. Berdasarkan Tabel 1, field yang dijadikan parameter target (class), yaitu field 'CEK' yang berisi nilai parameter AKTIF dan TIDAK-AKTIF. Nilai parameter AKTIF berarti nasabah masih mempunyai angsuran kredit kepada bank, sedangkan TIDAK-AKTIF berarti nasabah sudah tidak mempunyai angsuran kredit kepada bank lagi (sudah melunasi pembayaran kreditnya).

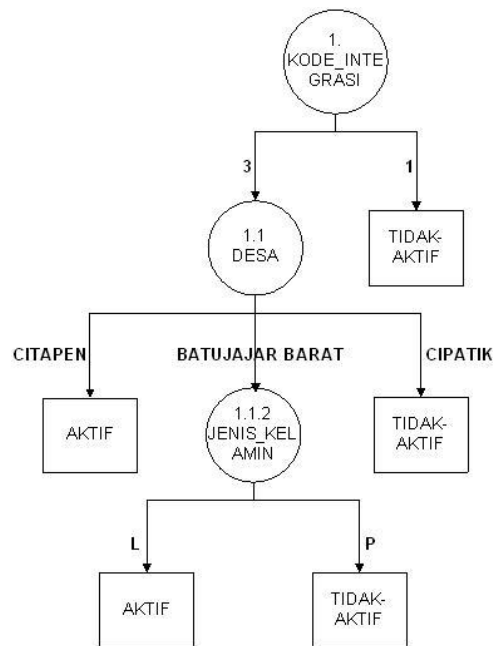
Berdasarkan records dari data pada field 'CEK' tersebut, maka teknik classification bisa diterapkan untuk melakukan data mining pada data tersebut. Adapun kolom yang bisa diambil sebagai parameter-parameter pembentuk pohon keputusannya adalah :

- JENIS_KELAMIN
- UMUR
- DESA
- KECAMATAN
- STATUS_MARITAL
- NILAI_PINJAMAN
- JML_ANGSURAN
- KODE_INTEGRASI

Contoh pohon keputusan yang terbentuk dengan menggunakan algoritma C4.5 dapat dilihat pada

gambar 2. Setelah pohon keputusan terbentuk secara utuh, maka dibuat aturan prediksi berdasarkan hasil pembentukan pohon keputusan tersebut. Adapun rules (aturan) prediksi yang dapat diperoleh dari pembentukan pohon keputusan pada contoh tersebut adalah sebagai berikut :

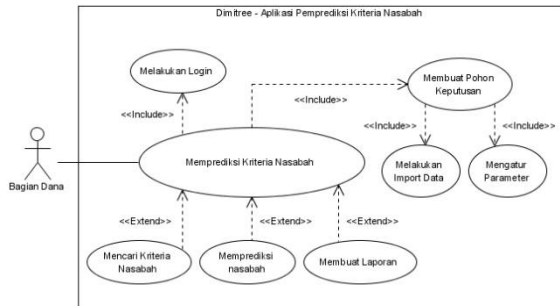
- Jika Pekerjaannya di bidang Perdagangan Dan DESA CITAPEN Diprediksikan Berpotensi
- Jika Pekerjaannya di bidang Perdagangan Dan DESA BATUJAJAR BARAT Dan Jenis Kelaminnya Laki-Laki Diprediksikan Berpotensi
- Jika Pekerjaannya di bidang Perdagangan Dan DESA CIPATIK Diprediksikan Tidak Berpotensi
- Jika Pekerjaannya di bidang Perdagangan Dan DESA BATUJAJAR BARAT Dan Jenis Kelaminnya Perempuan Diprediksikan Tidak Berpotensi
- Jika Pekerjaannya di bidang Pertanian Diprediksikan Tidak Berpotensi



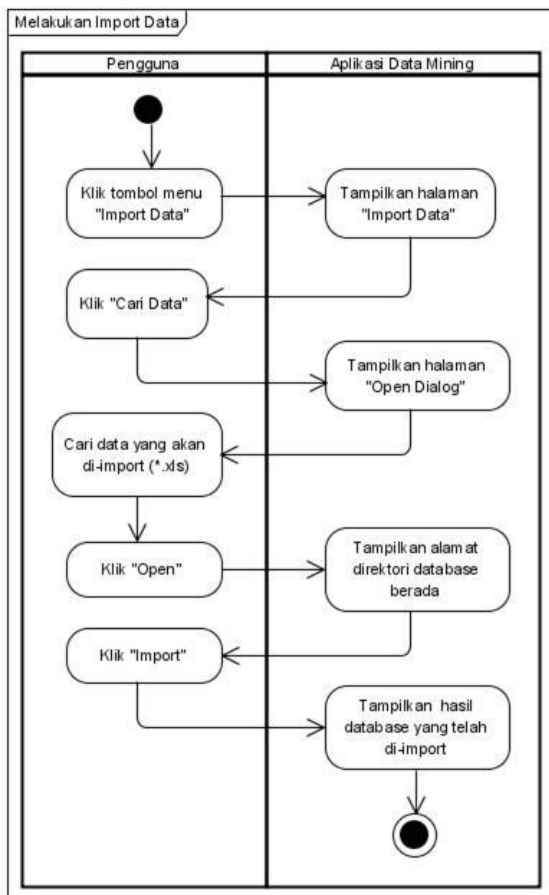
Gambar 2. Pohon Keputusan Node 1.1.2

Gambar 3 menunjukkan use case aplikasi data mining yang dibangun sehingga tampak jelas interaksi antara pengguna dengan aplikasi data mining. Aktor yang terkait dengan aplikasi data mining ini adalah Bagian Dana. Aktor ini mempunyai wewenang sepenuhnya untuk menangani aplikasi seperti melakukan login, import data, mengatur parameter, memproses pohon keputusan, melakukan prediksi, menguji kriteria nasabah, dan membuat laporan. Interaksi antara pengguna dengan aplikasi data mining yang terjadi

pada setiap aktivitas di atas, akan diperjelas dengan menggunakan *activity diagram*. Sebagai contoh dapat dilihat Gambar 4 yang menunjukkan *activity diagram* “Mengimport Data”.



Gambar 3. Use Case Diagram Aplikasi Data Mining

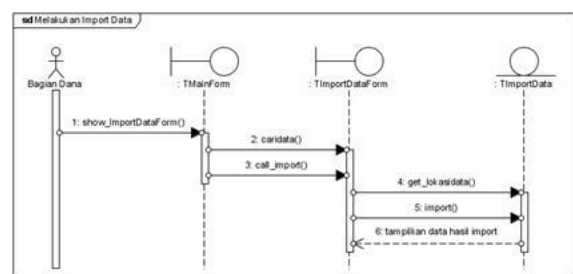


Gambar 4. Activity Diagram Mengimport Data

Berdasarkan skenario *use case*, akan dibentuk diagram *sequence* untuk kasus normal dari setiap skenario *use case*. Tabel 2 menunjukkan skenario *use case* melakukan *import data* dan Gambar 5 menunjukkan *sequence diagram*nya. Sedangkan Gambar 6 menunjukkan *class diagram* sebagai hasil realisasi dari *use case diagram*.

Tabel 2
Skenario *Use Case* Melakukan *Import Data*

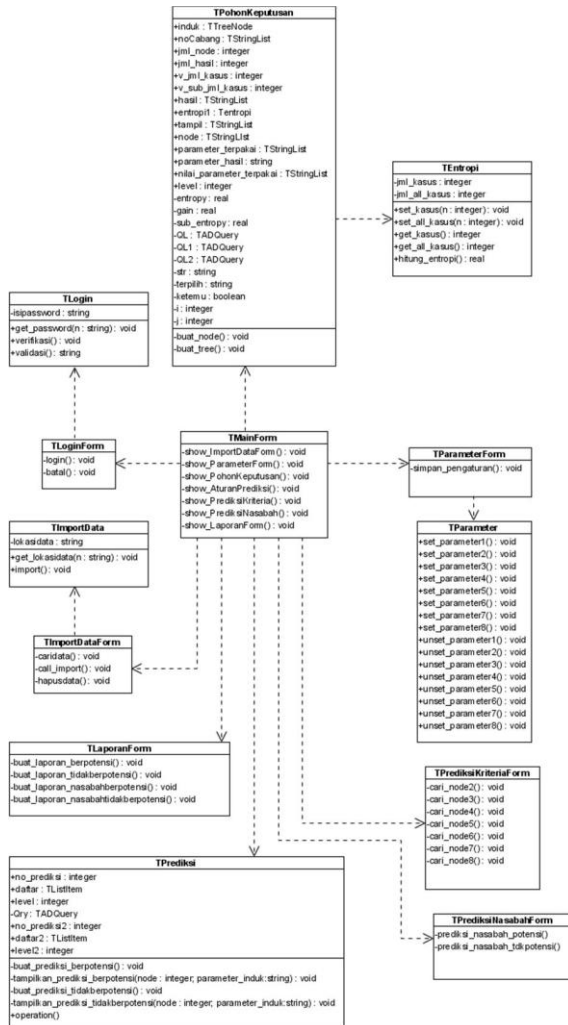
| Identifikasi | |
|---|---|
| Nomor | UC-02 |
| Nama | Melakukan <i>import Data</i> |
| Tujuan | Mengimport data yang akan diolah atau dianalisis |
| Identifikasi | |
| Deskripsi | Sistem membaca lokasi alamat data yang diinputkan pengguna, melakukan proses <i>import</i> data setelah aktor mengklik tombol “ <i>import</i> ”, kemudian menampilkan hasil data yang telah selesai di impor ke dalam aplikasi. |
| Aktor | Bagian Dana |
| Skenario Utama | |
| Kondisi awal | Pengguna sudah terautentikasi dan berada pada halaman utama aplikasi <i>data mining</i> |
| Aksi Aktor | Reaksi Sistem |
| 1. Mengklik tombol menu “ <i>Import Data</i> ” | |
| | 2. Menampilkan halaman <i>Import Data</i> |
| 3. Mengklik tombol “ <i>Browse</i> ” | |
| | 4. Menampilkan halaman <i>Open Dialog</i> |
| 5. Mencari lokasi <i>database</i> nasabah dalam format <i>Microsoft Excel (*.xls)</i> | |
| 6. Setelah ditemukan kemudian klik tombol “ <i>Open</i> ” | |
| | 7. Menampilkan lokasi alamat <i>database</i> |
| 8. Klik tombol “ <i>Import Data</i> ” untuk memulai proses <i>import data</i> | |
| | 9. Mengimport data |
| | 10. Menampilkan hasil data yang telah diimport ke dalam sistem. |
| Skenario Alternatif – Lokasi data belum ditentukan | |
| Aksi Aktor | Reaksi Sistem |
| 1. Mengklik tombol “ <i>Import Data</i> ” | |
| | 2. Tombol “ <i>Import Data</i> ” tidak aktif atau <i>disabled</i> . |
| 2. Mencari lokasi <i>database</i> nasabah dalam format <i>Microsoft Excel (*.xls)</i> . | |
| Kondisi akhir | Data berhasil diimport |



Gambar 5. Sequence Diagram Melakukan *Import Data*

DAFTAR PUSTAKA

- [1] Agushinta, D., Irfan, M. (2008), Perancangan Aplikasi *Data Mining* Untuk Memprediksi Permintaan *Customer* Pada Perusahaan Persewaan Mobil, Prosiding Seminar Ilmiah Nasional Komputer dan Sistem Intelijen (KOMMIT 2008), Depok, Universitas Gunadarma, 207-208
- [2] Berry, M.J.A., (2004), *Data Mining Techniques*, John Wiley & Sons.
- [3] Giudici, P., (2003), *Applied Data Mining Statistical Methods for Business and Industry*, England : John Wiley & Sons, Ltd.
- [4] Han, J., Micheline, K. (2006), *Data Mining Concepts and Techniques, Second Edition*. California : Morgan Kauffman Publishers.
- [5] Khairina, I.K. (____), Penggunaan Pohon Keputusan untuk *Data Mining*, ITB,1-3,<http://www.informatika.org/~rinaldi/Matdis/2008-2009/Makalah2008/Makalah0809-005.pdf>



Gambar 6. Class Diagram Aplikasi Data Mining

4. KESIMPULAN DAN SARAN

Setelah melakukan analisis, perancangan, implementasi beserta pengujian yang telah dilakukan, maka dapat disimpulkan bahwa aplikasi yang dibangun dapat membantu Bagian Dana dalam menganalisis data nasabah untuk menentukan target pemasaran kredit sehingga diharapkan biaya operasional *marketing* perbankan dapat ditekan seminimal mungkin. Jadi perancangan aplikasi data mining ini telah sesuai dengan prosedur dan tujuan yang diharapkan. Adapun saran-saran untuk pengembangan aplikasi ini lebih lanjut adalah :

1. Pembersihan data diharapkan bisa terintegrasi dalam aplikasi.
2. Adanya grafik yang mempresentasikan hasil analisis.
3. Format masukan data tidak hanya berupa *Microsoft Excel* saja tetapi bisa *Microsoft Access*, *SQL server*, dan lain sebagainya.